3D Shape Modelling Using a Multi-scale Surface Model

Xinquan Shen and Michael Spann
School of Electronic and Electrical Engineering
The University of Birmingham
Birmingham, B15 2TT, UK

Abstract

This paper presents a method for extracting the 3D shape of objects from 3D gray level images. The object surface is assumed to be described by a bicubic B-spline surface. The detection of the object shape is achieved using a hierarchical approach in which the object surface is successively estimated using the spline model at increased scales. At each scale, the model is iteratively fitted to a set of measurement points which are believed to be on the object surface. The measurement points are detected by searching for voxels with the strongest edge response along the direction normal to model surface. Searching for measurement points is controlled within a search radius determined based on the edge strength at the points on the current model surface. The result is a parametric surface approximating the object surface to be detected.

1 Introduction

An important task in the analysis of 3D images is to compute the 3D geometric structure of objects in the image. This information can be obtained from the 3D shape of the object which may be extracted by detecting all the surface elements belonging to the same object and integrating them into a shape model mathematically describing the object surface.

Methods using deformable/active surface models [1, 2, 3, 4] detect the object surface through the deformation of the model by imposing external forces to drive the model towards the object surface. The deformation of the model is determined by internal forces of the model and external forces acted on the model, and is computed by minimising the energy of the model. To cope with the noise effect in the image, a regularisation term acting as the internal energy of the model is often used to restrain the model to be deformed in some expected way (e.g. maintaining the smoothness of the model surface [1]). The stiffness of the model is uniquely determined by the internal energy of the model. In practice, the stiffness coefficients are often chosen empirically based on the potential object shape, noise level, boundary irregularities of the object, etc. Unsuitable stiffness coefficients may lead to incorrect results. Since only a local minimisation of the model energy can be guaranteed, the methods potentially need to arrange the initial model near the object to be detected, although an external "inflation"/"deflation" energy [3] could be devised to compensate for this limitation.

In this paper, we present a method for extracting the 3D shape of the object from 3D gray level images. A bicubic B-spline surface is used as a surface model. The object surface is estimated by iteratively fitting the spline model to a set of measurement points which are believed to be located on the boundary of the object. At each iteration, measurement points are detected by searching, within the respective search radii, the voxels with strongest edge response along directions normal to the model surface. The search radius associated with each point on the model is automatically controlled by the edge strength at the point. The surface fitting is performed in a hierarchical way by exploiting a multi-scale representation of the model obtained by using a variable number of control points defining the model surface. First, the model at the lowest scale is used to obtain a coarse estimation of the object surface, leading to a quick location of the object surface and an immunity to noise. The estimation is then gradually refined by using the model at increased scales. The model surface is automatically maintained as a C^2 continuous surface during surface fitting, without using any regularisation. The estimated surface is thus a C^2 continuous parametric surface approximating the object surface to be detected.

2 B-spline shape model

The object surface $\boldsymbol{S}(u,v)$ is parametrically represented by a uniform bicubic B-spline surface expressed as follows [5]:

$$\begin{aligned}\boldsymbol{S}^{(k)}(u,v) &= [x(u,v)\,y(u,v)\,z(u,v)]^T \\ &= \sum_{i=0}^{M_k-1}\sum_{j=0}^{N_k-1} \boldsymbol{X}_{ij}^{(k)} B_i(u) B_j(v) \qquad (1)\end{aligned}$$

where, $B_i(u)$ and $B_j(v)$ are B-spline basis functions of order 4 defined on the parameters u ($0 \le u \le M_k$) and v ($0 \le v \le N_k$) respectively, $\boldsymbol{X}_{ij}^{(k)} = [x_{ij}\,y_{ij}\,z_{ij}]^T$ are the control points defining the surface, and $k \geq 0$ indexes to a level in the hierarchy of the model with different scales. M_k and N_k determine the number of control points used for defining the surface and therefore control the scale of the model. At a given scale k, M_k and N_k are chosen as $M_k = m + ck$ and $N_k = n + ck$, where m and n determine the number of control points used for the model at the coarsest scale, and c is a constant specifying the scale increment.

Given a set of measurement points $\boldsymbol{S}'^{(k)}(u_s, v_t)$, $s = 0, 1, \ldots, P_k$, and $t = 0, 1, \ldots, Q_k$, we estimate the control points of the model by minimising the following error:

$$\begin{aligned}\text{Error} \;=\; & \sum_{s=0}^{P_k-1} \sum_{t=0}^{Q_k-1} \Big[\boldsymbol{S}'^{(k)}(u_s, v_t) - \\ & \sum_{i=0}^{M_k-1} \sum_{j=0}^{N_k-1} \boldsymbol{X}_{ij}^{(k)} B_i(u_s) B_j(v_t) \Big]^2 \qquad (2)\end{aligned}$$

P_k and Q_k determine the number of measurement points used for estimating the model, and are chosen to satisfy the conditions $P_k = p + ck$, $p \geq m$, and $Q_k = q + ck$, $q \geq n$. By vanishing the first order derivatives with respect to the $\boldsymbol{X}_{ij}^{(k)}$, the control points can be estimated by solving the following matrix equations [5]:

$$\mathcal{X}^{(k)} = \left(\mathcal{B}^{(k)T} \mathcal{B}^{(k)} \right)^{-1} \mathcal{B}^{(k)T} \mathcal{S}^{(k)} \qquad (3)$$

where, $\mathcal{B}^{(k)}$ is a $P_k Q_k \times M_k N_k$ matrix computed from the basis functions, $\mathcal{S}^{(k)}$ is a $P_k Q_k \times 3$ matrix consisting of the coordinates of all measurement points, and $\mathcal{X}^{(k)}$ is a $M_k N_k \times 3$ matrix consisting of the coordinates of the control points.

The surface defined in (1) is C^2 continuous. Thus, a C^2 smoothed estimation is always ensured, without explicitly using any constraint.

3 Extracting the object surface

To estimate the object surface, a set of measurement points which are believed to be located on the boundary of the object region is needed. It is generally true that voxels on the object boundary are edges. Thus, we detect the measurement points by searching voxels with strong edge responses. Given an initial model at a scale k, a set of $P_k \times Q_k$ sample points $\boldsymbol{S}^{(k)}(u_s, v_t)$ are first computed by uniformly sampling the model surface. At each sample point $\boldsymbol{S}^{(k)}(u_s, v_t)$, a measurement point $\boldsymbol{S}'^{(k)}(u_s, v_t)$ is then detected by searching, within a neighbourhood of $\boldsymbol{S}^{(k)}(u_s, v_t)$, for a voxel with the strongest edge response along the direction of the normal[1] at the point $\boldsymbol{S}^{(k)}(u_s, v_t)$. To ensure a valid estimation with the uniform B-splines, the obtained measurement points are adjusted by applying linear interpolation so that they are equally spaced in u_s and v_t.

[1] Depending on the initial state of the model, the search direction can be along the positive direction, negative direction, or both directions of the normal.

The measurement points so obtained will attract the model surface towards positions having local maximal edge responses. However, an unsuitable neighbourhood for searching a measurement point may lead to a voxel either belonging to another object or belonging to the background. Using a fixed neighbourhood size for each sample point makes the problem even worse. A solution to this is to choose the neighbourhood based on the measurement covariance [6]. In this work, we associate with each sample point $\boldsymbol{S}^{(k)}(u_s, v_t)$ a search radius $r^{(k)}(u_s, v_t)$ based on the confidence for using the $\boldsymbol{S}^{(k)}(u_s, v_t)$ as the measurement point. In general, a strong edge response at the point $\boldsymbol{S}^{(k)}(u_s, v_t)$ suggests that $\boldsymbol{S}^{(k)}(u_s, v_t)$ itself is very likely a true measurement point and a small search radius should be used, while a weak edge response indicates that the true measurement point is somewhere away from the current point and therefore a large search radius is needed. The search radius $r^{(k)}(u_s, v_t)$ is thus defined as:

$$r^{(k)}(u_s, v_t) = 1 + d \frac{\delta_h - \varphi(\boldsymbol{S}^{(k)}(u_s, v_t))}{(\delta_h - \delta_l)} \qquad (4)$$

where, $\varphi(\boldsymbol{S}^{(k)}(u_s, v_t))$ is determined as follows:

$$\varphi(\boldsymbol{S}^{(k)}(u_s, v_t)) = \begin{cases} \delta_l & \text{if } \Phi(\boldsymbol{S}^{(k)}(u_s, v_t)) < \delta_l \\ \delta_h & \text{if } \Phi(\boldsymbol{S}^{(k)}(u_s, v_t)) > \delta_h \\ \Phi(\boldsymbol{S}^{(k)}(u_s, v_t)) & \text{otherwise} \end{cases}$$

where, $\Phi(\boldsymbol{S}^{(k)}(u_s, v_t)) > 0$ is the edge response at the point $\boldsymbol{S}^{(k)}(u_s, v_t)$, δ_l and δ_h are two thresholds for considering the $\boldsymbol{S}^{(k)}(u_s, v_t)$ as non-edge point and edge point respectively, and d is a constant controlling the largest possible search radius. $r^{(k)}(u_s, v_t)$ is thus within the range $[1, d+1]$.

An iterative procedure is necessary to gradually attract the model surface towards the object surface. At each scale, the model surface is iteratively estimated using

$$\mathcal{X}_{\lambda+1}^{(k)} = \left(\mathcal{B}^{(k)T} \mathcal{B}^{(k)} \right)^{-1} \mathcal{B}^{(k)T} \mathcal{S}_{\lambda}^{(k)}, \quad \lambda \geq 0 \qquad (5)$$

where, $\mathcal{S}_{\lambda}^{(k)}$ contains the measurement points obtained based on the estimation $\mathcal{X}_{\lambda}^{(k)}$. The iteration terminates when $\|\mathcal{S}_{\lambda+1}^{(k)} - \mathcal{S}_{\lambda}^{(k)}\| < \epsilon$. It should be noted that, for a fixed scale of the model, the matrix $\mathcal{B}$ (and therefore $(\mathcal{B}^{(k)T} \mathcal{B}^{(k)})^{-1} \mathcal{B}^{(k)T}$) is fixed and only need to be computed once.

The effect of the measurement points for estimating the model surface is dominated by the number of control points defining the model. A model with a small number of control points is inclined to have a stiff surface in response to the measurement points, leading to an over-smoothed estimation. As the number of control points increases, the "stiffness" in response to the measurement points is reduced, encouraging an estimation passing through the measurement points. Consequently, we apply the surface fitting in a hierarchical way by exploiting the multi-scale nature of the model. First, the model at the lowest scale ($k = 0$) is used to achieve a rough (yet quick) estimation of the object surface. This will prevent the model surface from being tied to positions where the image is contaminated by the noise and produces strong edge responses. The result is then refined by increasing the scale of the model up to the desired scale level ($k = K$). At each scale, the object surface is estimated using the iterative surface fitting (5). Thus, local details of the object shape will be gradually extracted as the model scale increases. Given a model at a scale k, $\boldsymbol{S}^{(k)}(u, v)$, the initial model at the scale $k + 1$ can be easily produced by computing $P_{k+1} \times Q_{k+1}$ uniformly distributed sample points from the $\boldsymbol{S}^{(k)}(u, v)$ and estimating $M_{k+1} \times N_{k+1}$ control points using these sample points. Therefore, the hierarchical surface fitting is triggered by giving an initial model at the lowest scale.

4 Experimental results

Experiments have been applied to both synthetic and real 3D images. The 3D edge operator proposed by Zucker and Hummel [7] is used in the detection of measurement points. Fig. 1 gives the result of an experiment applied to a $64 \times 64 \times 64$ synthetic image (showing 5 XY slices overlapped with the model surface). The image with additive white Gaussian noise contains a $32 \times 32 \times 32$ block. Fig. 1a shows, from the top to bottom, the initial model (a sphere with radius 3), the result using a single scale model with $M = 9$, $N = 10$, $P = 15$, and $Q = 16$, the result using a single scale model with $M = 21$, $N = 22$, $P = 27$, and $Q = 28$, and the result using a multi-scale model (with $M_0 = 9$, $N_0 = 10$, $P_0 = 15$, $Q_0 = 16$, $c = 2$, and $K = 6$). Fig. 1b gives a 3D visualisation of the estimated surface shown in the last row of Fig. 1a. It can be seen that the boundary of the object is more accurately detected by using the multi-scale model. Fig. 2 shows an experiment applied to a 3D image block extracted from a $256 \times 256 \times 180$ MRI brain image. Only 3 slices containing the object region of interest are shown. Fig. 2a shows, from the top to bottom, the initial model (a sphere), the result using a single scale model ($M = 9$, $N = 16$, $P = 15$, and $Q = 16$), and the result using a multi-scale model ($M_0 = 9$, $N_0 = 16$, $P_0 = 15$, $Q_0 = 16$, $c = 2$, and $K = 6$). The 3D visualisation of the estimated surface shown in the last row of Fig. 2a. is given in Fig. 2b. In this case, a single scale model at a high scale produces an even worse result than a single scale model at a low scale. Clearly, local details of the object shape are correctly extracted with the multi-scale model.

5 Conclusion

We have presented a method for detecting the object surface from the 3D gray level images. Modelling the object surface with a bicubic B-spline surface, the extraction of the object surface is achieved through a hierarchical surface estimation in which the model at different scales is successively fitted to a set of measurement points detected by searching the strongest edge response along the direction normal to the current model surface. The model is progressively refined to approach the object surface by detecting the measurement points within a search radius determined based on the edge strength at the current model surface.

Acknowledgement

This work was funded by the EPSRC IMV initiative (grant of GR/K 95987) and carried out in collaboration with Shell KSEPL.

References

[1] D. Terzopoulos, A. Witkin, and M. Kass. Constraints on Deformable Models: Recovering 3D Shape and Nonrigid Motion. *Artificial Intelligence*, 36(1):91–123, 1988.

[2] I. Cohen, L. D. Cohen, and N. Ayache. Using Deformable Surfaces to Segment 3-D Images and Infer Differential Structures. *Computer Vision, Graphics and Image Processing: Image Understanding*, 56(2):242–263, September 1992.

[3] T. McInerney and D. Terzopoulos. Medical Images Segmentation Using Topologically Adaptable Surfaces. In *Proc. CVMed'97*, Grenoble, France, March 1997.

[4] V. Caselles, R. Kimmmel, G. Sapiro, and C. Sbert. Minimal Surfaces Based Object Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-19(4):394–398, April 1997.

[5] R. H. Bartels, J. C. Beatty, and B. A. Barsky. *An Introduction to Splines for use in Computer Graphics and Geometric Modelling*. Morgan Kaufmann, 1987.

[6] A. Blake, R. Curwen, and A. Zisserman. A Framework for Spatio-temporal Control in the Tracking of Visual Contours. *International Journal of Computer Vision*, 11(2):127–145, 1993.

[7] S. W. Zucker and R. A. Hummel. A Three-Dimensional Edge Operator. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 3(3):324–331, May 1981.

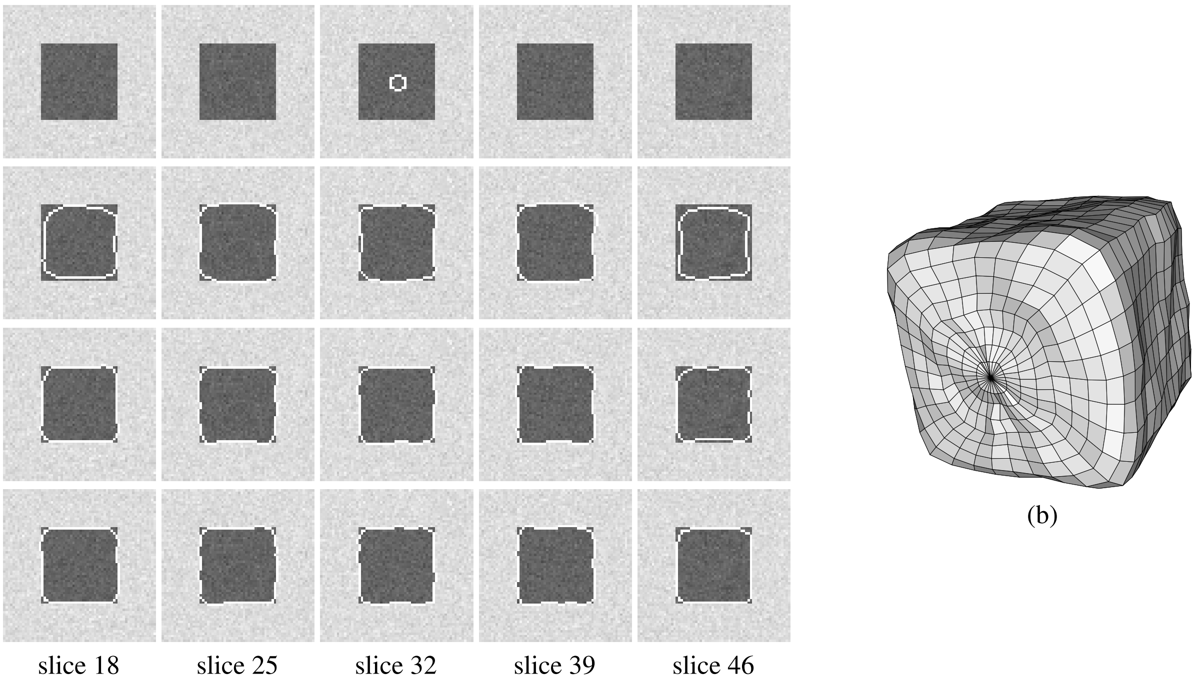

Figure 1: Experiments on a synthetic image. (a) from the top to bottom, the initial model (a sphere), the result using a single scale model with $M=9$, $N=10$, $P=15$, and $Q=16$, the result using a single scale model with $M=21$, $N=22$, $P=27$, and $Q=28$, and the result using a multi-scale model (with $M_0=9$, $N_0=10$, $P_0=15$, $Q_0=16$, $c=2$, and $K=6$). (b) a 3D visualisation of the estimated surface shown in the last row in (a).

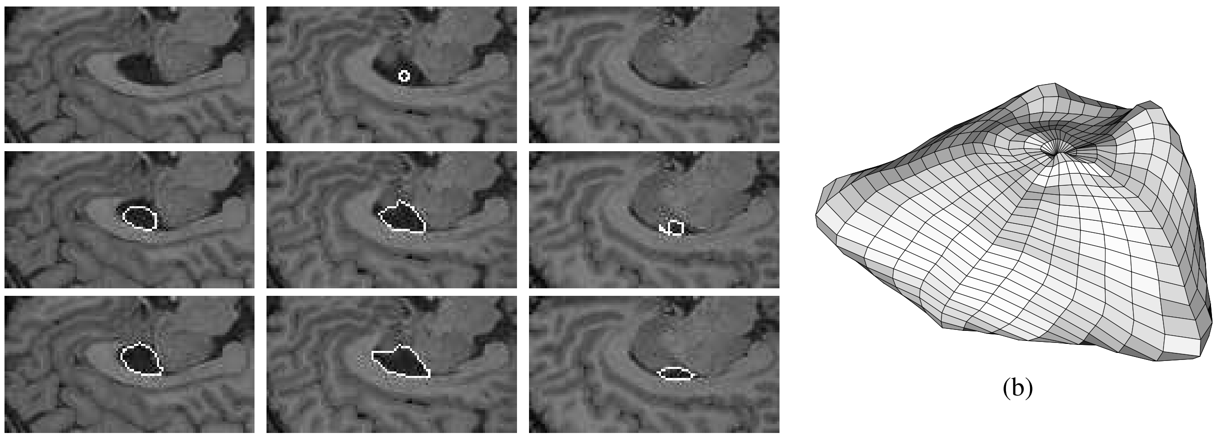

Figure 2: Experiments on a MRI brain image. (a) from the top to bottom, the initial model (a sphere), the result using a single scale model ($M=9$, $N=16$, $P=15$, and $Q=16$), and the result using a multi-scale model ($M_0=9$, $N_0=16$, $P_0=15$, $Q_0=16$, $c=2$, and $K=6$). (b) a 3D visualisation of the estimated surface shown in the last row in (a).